

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/582,002  
Source: IFWP  
Date Processed by STIC: 06/19/2006

# ***ENTERED***



IFWP

## RAW SEQUENCE LISTING

DATE: 06/19/2006

PATENT APPLICATION: US/10/582,002

TIME: 11:36:29

Input Set : A:\Q95278 Sequence Listing.txt

Output Set: N:\CRF4\06192006\J582002.raw

```

3 <110> APPLICANT: MEIJI SEIKA KAISHA, LTD.
4     WATANABE, Manabu
5     YANAI, Koji
6     TSUYUKI, Yumiko
8 <120> TITLE OF INVENTION: Surfactant tolerant cellulase and method for
modification thereto
10 <130> FILE REFERENCE: Q95278
C--> 12 <140> CURRENT APPLICATION NUMBER: US/10/582,002
C--> 12 <141> CURRENT FILING DATE: 2006-06-07
12 <150> PRIOR APPLICATION NUMBER: PCT/JP2004/018184
13 <151> PRIOR FILING DATE: 2004-12-07
15 <150> PRIOR APPLICATION NUMBER: JP 2003-409692
16 <151> PRIOR FILING DATE: 2003-12-08
18 <160> NUMBER OF SEQ ID NOS: 48
20 <170> SOFTWARE: PatentIn version 3.1
22 <210> SEQ ID NO: 1
23 <211> LENGTH: 870
24 <212> TYPE: DNA
25 <213> ORGANISM: Humicola insolens MN200-1
27 <220> FEATURE:
28 <221> NAME/KEY: CDS
29 <222> LOCATION: (1)..(870)
31 <220> FEATURE:
32 <221> NAME/KEY: misc_feature
33 <222> LOCATION: (1)..(3)
34 <223> OTHER INFORMATION: Pyroglutamic acid
37 <220> FEATURE:
38 <221> NAME/KEY: source
39 <222> LOCATION: (16)..(870)
40 <223> OTHER INFORMATION: Humicola insolens MN200-1
43 <400> SEQUENCE: 1
44 cag aac tgt gga tcc gct gat ggc aag tcc acc cgc tac tgg gac tgc      48
45 Gln Asn Cys Gly Ser Ala Asp Gly Lys Ser Thr Arg Tyr Trp Asp Cys
46 1           5           10           15
48 tgc aag cct tcg tgc ggc tgg gcc aag aag gct ccc gtg aac cag cct      96
49 Cys Lys Pro Ser Cys Gly Trp Ala Lys Lys Ala Pro Val Asn Gln Pro
50           20           25           30
52 gtc ttc tcc tgc aac gcc aac ttc cag cgt ctc act gac ttc gac gcc      144
53 Val Phe Ser Cys Asn Ala Asn Phe Gln Arg Leu Thr Asp Phe Asp Ala
54           35           40           45
56 aag tcc ggc tgc gag ccg ggc ggt gtc gcc tac tcg tgc gcc gac cag      192
57 Lys Ser Gly Cys Glu Pro Gly Gly Val Ala Tyr Ser Cys Ala Asp Gln
58           50           55           60
60 acc cca tgg gct gtg aac gac gac ttc gcg ttc ggt ttt gct gcc acc      240

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/582,002

DATE: 06/19/2006

TIME: 11:36:29

Input Set : A:\Q95278 Sequence Listing.txt

Output Set: N:\CRF4\06192006\J582002.raw

```

61 Thr Pro Trp Ala Val Asn Asp Asp Phe Ala Phe Gly Phe Ala Ala Thr
62 65 70 75 80
64 tct att gcc ggc agc aat gag gcg ggc tgg tgc tgc gcc tgc tac gag 288
65 Ser Ile Ala Gly Ser Asn Glu Ala Gly Trp Cys Cys Ala Cys Tyr Glu
66 85 90 95
68 ctc acc ttc aca tcc ggt cct gtt gct ggc aag aag atg gtc gtc cag 336
69 Leu Thr Phe Thr Ser Gly Pro Val Ala Gly Lys Lys Met Val Val Gln
70 100 105 110
72 tcc acc agc act ggc ggt gat ctt ggc agc aac cac ttc gat ctc aac 384
73 Ser Thr Ser Thr Gly Gly Asp Leu Gly Ser Asn His Phe Asp Leu Asn
74 115 120 125
76 atc ccc ggc ggc ggc gtc ggc atc ttc gac gga tgc act ccc cag ttc 432
77 Ile Pro Gly Gly Gly Val Gly Ile Phe Asp Gly Cys Thr Pro Gln Phe
78 130 135 140
80 ggc ggt ctg ccc ggc cag cgc tac ggc ggc atc tcg tcc cgc aac gag 480
81 Gly Gly Leu Pro Gly Gln Arg Tyr Gly Gly Ile Ser Ser Arg Asn Glu
82 145 150 155 160
84 tgc gat cgg ttc ccc gac gcc ctc aag ccc ggc tgc tac tgg cgc ttc 528
85 Cys Asp Arg Phe Pro Asp Ala Leu Lys Pro Gly Cys Tyr Trp Arg Phe
86 165 170 175
88 gac tgg ttc aag aac gcc gac aac ccg agc ttc agc ttc cgt cag gtc 576
89 Asp Trp Phe Lys Asn Ala Asp Asn Pro Ser Phe Ser Phe Arg Gln Val
90 180 185 190
92 caa tgc cca gcc gag ctc gtc gct cgc acc gga tgc cgc cgc aac gac 624
93 Gln Cys Pro Ala Glu Leu Val Ala Arg Thr Gly Cys Arg Arg Asn Asp
94 195 200 205
96 gac ggc aac ttc cct gcc gtc cag atc ccc tcc agc agc acc agc tct 672
97 Asp Gly Asn Phe Pro Ala Val Gln Ile Pro Ser Ser Ser Thr Ser Ser
98 210 215 220
100 ccg gtc ggc cag cct acc agt acc agc acc acc tcc acc tcc acc acc 720
101 Pro Val Gly Gln Pro Thr Ser Thr Ser Thr Ser Thr Ser Thr Thr
102 225 230 235 240
104 tcg agc ccg ccc gtc cag cct acg act ccc agc ggc tgc act gct gag 768
105 Ser Ser Pro Pro Val Gln Pro Thr Thr Pro Ser Gly Cys Thr Ala Glu
106 245 250 255
108 agg tgg gct cag tgc ggc ggc aat ggc tgg agc ggc tgc acc acc tgc 816
109 Arg Trp Ala Gln Cys Gly Gly Asn Gly Trp Ser Gly Cys Thr Thr Cys
110 260 265 270
112 gtc gct ggc agc acc tgc acg aag att aat gac tgg tac cat cag tgc 864
113 Val Ala Gly Ser Thr Cys Thr Lys Ile Asn Asp Trp Tyr His Gln Cys
114 275 280 285
116 ctg taa 870
117 Leu
121 <210> SEQ ID NO: 2
122 <211> LENGTH: 289
123 <212> TYPE: PRT
124 <213> ORGANISM: Humicola insolens MN200-1
126 <220> FEATURE:
127 <221> NAME/KEY: misc_feature

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/582,002

DATE: 06/19/2006

TIME: 11:36:29

Input Set : A:\Q95278 Sequence Listing.txt

Output Set: N:\CRF4\06192006\J582002.raw

```

128 <222> LOCATION: (1)..(3)
129 <223> OTHER INFORMATION: Pyroglutamic acid
131 <400> SEQUENCE: 2
133 Gln Asn Cys Gly Ser Ala Asp Gly Lys Ser Thr Arg Tyr Trp Asp Cys
134 1          5          10          15
137 Cys Lys Pro Ser Cys Gly Trp Ala Lys Lys Ala Pro Val Asn Gln Pro
138          20          25          30
141 Val Phe Ser Cys Asn Ala Asn Phe Gln Arg Leu Thr Asp Phe Asp Ala
142          35          40          45
145 Lys Ser Gly Cys Glu Pro Gly Gly Val Ala Tyr Ser Cys Ala Asp Gln
146          50          55          60
149 Thr Pro Trp Ala Val Asn Asp Asp Phe Ala Phe Gly Phe Ala Ala Thr
150 65          70          75          80
153 Ser Ile Ala Gly Ser Asn Glu Ala Gly Trp Cys Cys Ala Cys Tyr Glu
154          85          90          95
157 Leu Thr Phe Thr Ser Gly Pro Val Ala Gly Lys Lys Met Val Val Gln
158          100         105         110
161 Ser Thr Ser Thr Gly Gly Asp Leu Gly Ser Asn His Phe Asp Leu Asn
162          115         120         125
165 Ile Pro Gly Gly Gly Val Gly Ile Phe Asp Gly Cys Thr Pro Gln Phe
166          130         135         140
169 Gly Gly Leu Pro Gly Gln Arg Tyr Gly Gly Ile Ser Ser Arg Asn Glu
170 145         150         155         160
173 Cys Asp Arg Phe Pro Asp Ala Leu Lys Pro Gly Cys Tyr Trp Arg Phe
174          165         170         175
177 Asp Trp Phe Lys Asn Ala Asp Asn Pro Ser Phe Ser Phe Arg Gln Val
178          180         185         190
181 Gln Cys Pro Ala Glu Leu Val Ala Arg Thr Gly Cys Arg Arg Asn Asp
182          195         200         205
185 Asp Gly Asn Phe Pro Ala Val Gln Ile Pro Ser Ser Ser Thr Ser Ser
186          210         215         220
189 Pro Val Gly Gln Pro Thr Ser Thr Ser Thr Thr Ser Thr Ser Thr Thr
190 225         230         235         240
193 Ser Ser Pro Pro Val Gln Pro Thr Thr Pro Ser Gly Cys Thr Ala Glu
194          245         250         255
197 Arg Trp Ala Gln Cys Gly Gly Asn Gly Trp Ser Gly Cys Thr Thr Cys
198          260         265         270
201 Val Ala Gly Ser Thr Cys Thr Lys Ile Asn Asp Trp Tyr His Gln Cys
202          275         280         285
205 Leu
209 <210> SEQ ID NO: 3
210 <211> LENGTH: 897
211 <212> TYPE: DNA
212 <213> ORGANISM: Staphylotrichum coccosporum IFO 31817
215 <220> FEATURE:
216 <221> NAME/KEY: CDS
217 <222> LOCATION: (1)..(897)
219 <220> FEATURE:
220 <221> NAME/KEY: misc_feature

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/582,002

DATE: 06/19/2006

TIME: 11:36:29

Input Set : A:\Q95278 Sequence Listing.txt

Output Set: N:\CRF4\06192006\J582002.raw

```

221 <222> LOCATION: (1)..(3)
222 <223> OTHER INFORMATION: Pyroglutamic acid
225 <220> FEATURE:
226 <221> NAME/KEY: source
227 <222> LOCATION: (13)..(897)
228 <223> OTHER INFORMATION: Staphylotrichum coccosporum IFO 31817
231 <400> SEQUENCE: 3
232 cag tgc gca tgc gat ggc aag tcc acc cgc tac tgg gac tgc tgc aag      48
233 Gln Ser Ala Cys Asp Gly Lys Ser Thr Arg Tyr Trp Asp Cys Cys Lys
234 1          5          10          15
236 cct tgc tgc tgc tgg ccc ggc aag gcc tgc gtg aac cag ccc gtc ttc      96
237 Pro Ser Cys Ser Trp Pro Gly Lys Ala Ser Val Asn Gln Pro Val Phe
238          20          25          30
240 gcc tgc agc gcc aac ttc cag cgc atc agc gac ccc aac gtc aag tgc      144
241 Ala Cys Ser Ala Asn Phe Gln Arg Ile Ser Asp Pro Asn Val Lys Ser
242          35          40          45
244 ggc tgc gac ggc ggc tcc gcc tac gcc tgc gcc gac cag acc ccg tgg      192
245 Gly Cys Asp Gly Gly Ser Ala Tyr Ala Cys Ala Asp Gln Thr Pro Trp
246          50          55          60
248 gcc gtc aac gac aac ttc tgc tac ggc ttc gcc gcc acg tcc atc tgc      240
249 Ala Val Asn Asp Asn Phe Ser Tyr Gly Phe Ala Ala Thr Ser Ile Ser
250 65          70          75          80
252 ggc ggc aac gag gcc tgc tgg tgc tgt ggc tgc tac gag ctg acc ttc      288
253 Gly Gly Asn Glu Ala Ser Trp Cys Cys Gly Cys Tyr Glu Leu Thr Phe
254          85          90          95
256 acc tgc ggc ccc gtc gct ggc aag acc atg gtt gtc cag tcc acc tgc      336
257 Thr Ser Gly Pro Val Ala Gly Lys Thr Met Val Val Gln Ser Thr Ser
258          100          105          110
260 acc ggc ggc gac ctc ggc acc aac cac ttc gac ctg gcc atg ccc ggt      384
261 Thr Gly Gly Asp Leu Gly Thr Asn His Phe Asp Leu Ala Met Pro Gly
262          115          120          125
264 ggt ggt gtc ggc atc ttc gac ggc tgc tgc ccc cag ttc ggc ggc ctc      432
265 Gly Gly Val Gly Ile Phe Asp Gly Cys Ser Pro Gln Phe Gly Gly Leu
266          130          135          140
268 gcc ggc gac cgc tac ggc ggc gtc tgc tgc cgc agc cag tgc gac tgc      480
269 Ala Gly Asp Arg Tyr Gly Gly Val Ser Ser Arg Ser Gln Cys Asp Ser
270 145          150          155          160
272 ttc ccc gcc gcc ctc aag ccc ggc tgc tac tgg cgc ttc gac tgg ttc      528
273 Phe Pro Ala Ala Leu Lys Pro Gly Cys Tyr Trp Arg Phe Asp Trp Phe
274          165          170          175
276 aag aac gcc gac aac ccg acc ttc acc ttc cgc cag gtc cag tgc ccg      576
277 Lys Asn Ala Asp Asn Pro Thr Phe Thr Phe Arg Gln Val Gln Cys Pro
278          180          185          190
280 tgc gag ctc gtc gcc cgc acc ggc tgc cgc cgc aac gac gac ggc aac      624
281 Ser Glu Leu Val Ala Arg Thr Gly Cys Arg Arg Asn Asp Asp Gly Asn
282          195          200          205
284 ttc ccc gtc ttc acc cct ccc tgc ggc ggt cag tcc tcc tgc tct tcc      672
285 Phe Pro Val Phe Thr Pro Pro Ser Gly Gly Gln Ser Ser Ser Ser Ser
286          210          215          220

```

## RAW SEQUENCE LISTING

DATE: 06/19/2006

PATENT APPLICATION: US/10/582,002

TIME: 11:36:29

Input Set : A:\Q95278 Sequence Listing.txt

Output Set: N:\CRF4\06192006\J582002.raw

```

288 tcc tcc agc agc gcc aag ccc acc tcc acc tcc acc tcg acc acc tcc      720
289 Ser Ser Ser Ser Ala Lys Pro Thr Ser Thr Ser Thr Ser Thr Thr Ser
290 225                230                235                240
292 acc aag gct acc tcc acc acc tcg acc gcc tcc agc cag acc tcg tcg      768
293 Thr Lys Ala Thr Ser Thr Thr Ser Thr Ala Ser Ser Gln Thr Ser Ser
294                245                250                255
296 tcc acc ggc ggc ggc tgc gcc gcc cag cgc tgg gcg cag tgc ggc ggc      816
297 Ser Thr Gly Gly Gly Cys Ala Ala Gln Arg Trp Ala Gln Cys Gly Gly
298                260                265                270
300 atc ggg ttc tcg ggc tgc acc acg tgc gtc agc ggc acc acc tgc aac      864
301 Ile Gly Phe Ser Gly Cys Thr Thr Cys Val Ser Gly Thr Thr Cys Asn
302                275                280                285
304 aag cag aac gac tgg tac tcg cag tgc ctt taa                        897
305 Lys Gln Asn Asp Trp Tyr Ser Gln Cys Leu
306 290                295
309 <210> SEQ ID NO: 4
310 <211> LENGTH: 298
311 <212> TYPE: PRT
312 <213> ORGANISM: Staphylotrichum coccosporum IFO 31817
314 <220> FEATURE:
315 <221> NAME/KEY: misc_feature
316 <222> LOCATION: (1)..(3)
317 <223> OTHER INFORMATION: Pyroglutamic acid
319 <400> SEQUENCE: 4
321 Gln Ser Ala Cys Asp Gly Lys Ser Thr Arg Tyr Trp Asp Cys Cys Lys
322 1                5                10                15
325 Pro Ser Cys Ser Trp Pro Gly Lys Ala Ser Val Asn Gln Pro Val Phe
326                20                25                30
329 Ala Cys Ser Ala Asn Phe Gln Arg Ile Ser Asp Pro Asn Val Lys Ser
330                35                40                45
333 Gly Cys Asp Gly Gly Ser Ala Tyr Ala Cys Ala Asp Gln Thr Pro Trp
334 50                55                60
337 Ala Val Asn Asp Asn Phe Ser Tyr Gly Phe Ala Ala Thr Ser Ile Ser
338 65                70                75                80
341 Gly Gly Asn Glu Ala Ser Trp Cys Cys Gly Cys Tyr Glu Leu Thr Phe
342                85                90                95
345 Thr Ser Gly Pro Val Ala Gly Lys Thr Met Val Val Gln Ser Thr Ser
346                100               105               110
349 Thr Gly Gly Asp Leu Gly Thr Asn His Phe Asp Leu Ala Met Pro Gly
350                115               120               125
353 Gly Gly Val Gly Ile Phe Asp Gly Cys Ser Pro Gln Phe Gly Gly Leu
354 130               135               140
357 Ala Gly Asp Arg Tyr Gly Gly Val Ser Ser Arg Ser Gln Cys Asp Ser
358 145               150               155               160
361 Phe Pro Ala Ala Leu Lys Pro Gly Cys Tyr Trp Arg Phe Asp Trp Phe
362                165                170                175
365 Lys Asn Ala Asp Asn Pro Thr Phe Thr Phe Arg Gln Val Gln Cys Pro
366                180                185                190
369 Ser Glu Leu Val Ala Arg Thr Gly Cys Arg Arg Asn Asp Asp Gly Asn

```

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/582,002

DATE: 06/19/2006

TIME: 11:36:30

Input Set : A:\Q95278 Sequence Listing.txt

Output Set: N:\CRF4\06192006\J582002.raw

L:12 M:270 C: Current Application Number differs, Replaced Current Application No

L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date